

Mapper

```
class MAPPER
```

```
  method INITIALIZE()
```

```
    Create empty temp_list
```

```
    Create empty temp_dict
```

```
    Set old_user_id to some non-positive integer (say 0)
```

```
  method MAP(line l)
```

```
    info <- split the line
```

```
    user_id <- info[0]
```

```
    movie_id <- info[1]
```

```
    rating <- info[2]
```

```
    if rating >= 4.0 do
```

```
      if old_user_id != user_id do
```

```
        sort temp_list
```

```
        for all movie_i in temp_list do
```

```
          for all movie_j > movie_i in temp_list do
```

```
            if movie_j in temp_dict do
```

```
              temp_dict[movie_j] <- temp_dict[movie_j] + 1
```

```
            else
```

```
              temp_dict[movie_j] <- 1
```

```
            EMIT(movie_i, movie_dict)
```

```
            movie_dict <- {}
```

```
        temp_list <- []
```

```
        old_user_id <- user_id
```

```
      else
```

```
        temp_list.append(movie_id)
```

Reducer

class REDUCER

method INITIALIZE()

Read movies.csv file and store the movie names into a global dictionary

where movie_id is key and movie_name is the value

set threshold to any value (say 1000)

method REDUCE(movie_i, list [dict1, dict2,])

for all dict in list **do**

for all movie_j in dict **do**

if movie_j in temp_dict **do**

temp_dict[movie_j] <- temp_dict[movie_j] + 1

else

temp_dict[movie_j] <- 1

for all movie_j in temp_dict **do**

if temp_dict[movie_j] > threshold **do**

EMIT(movie_i, movie_j, temp_dict[movie_j])